The Measurement and Characteristics of Professional Forecasters' Uncertainty

Gianna Boero, Jeremy Smith and Kenneth F. Wallis*

Department of Economics
University of Warwick
Coventry CV4 7AL, UK

First version, 20 April 2012

Abstract There is increased interest in extracting indicators of macroeconomic risk and uncertainty from forecast surveys that include questions about density forecasts. In this paper we consider several statistical issues that arise in the construction and interpretation of measures of uncertainty from such data, with application to the Bank of England Survey of External Forecasters. We find substantial heterogeneity of individual forecast uncertainty, and significant persistence in individual relative uncertainty. This is an individual characteristic akin to the individual optimism and pessimism already established among point forecasts, and which is seen to be present in the current data, in a bivariate sense with respect to forecasts of inflation and output growth. Whether disagreement among point forecasts is a useful indicator of uncertainty is seen to depend on the underlying macroeconomic environment. In an appendix, the main findings of the paper are replicated with data from the European Central Bank Survey of Professional Forecasters.

Acknowledgments We are grateful to Bank of England staff for assembling and helping to clean the survey dataset. Readers wishing to gain access to the data should write to the Publications Editor, Inflation Report and Bulletin Division, Bank of England, Threadneedle Street, London EC2R 8AH, UK. The ECB SPF dataset is available at http://www.ecb.int/stats/prices/indic/forecast/html/index.en.html .

***Corresponding author** K.F.Wallis@warwick.ac.uk

1. Introduction

Macroeconomic risk and uncertainty is an abiding preoccupation of policy makers, analysts and researchers, who make regular use of a wide range of indirect measures, indicators and proxy variables when assessing the current economic conjuncture and future economic prospects. In contrast to these proxy variables, a more direct assessment is provided by measures of uncertainty obtained from survey data, in particular from surveys of professional forecasters, which are receiving increasing attention. In a recent article Soderlind (2011), for example, shows how measures of inflation uncertainty based on survey data are a useful indicator of inflation risk premia, thereby helping to understand the evolution of 'break-even inflation' (the difference between nominal and real interest rates). Most of his empirical analysis is based on United States data, which include the US Survey of Professional Forecasters managed by the Federal Reserve Bank of Philadelphia, although he also considers two further countries that have survey data on inflation uncertainty. One is the United Kingdom, for which he uses a series constructed from the Bank of England Survey of External Forecasters in one of our earlier articles (Boero, Smith and Wallis, 2008); the second is the euro area, for which he uses the European Central Bank's Survey of Professional Forecasters. Soderlind (2011) also provides useful references to the related literature, some of which reappear below. Almost all of the empirical literature is based on the US data, whereas the empirical work in the present paper continues our analysis of the Bank of England dataset, which is updated annually; the appendix contains a brief extension to the ECB's survey.

The raw data on uncertainty in the three available quarterly surveys comprise the responses of individual members of a panel of forecasters to questions asking for their personal probabilities that the value of the variable of interest (inflation, output growth, …) in a specified future period will lie in each one of a number of preassigned intervals. Respondents thus supply density forecasts in the form of histograms. In this paper we consider several statistical issues that arise in the construction and interpretation of measures of uncertainty from these data, and discover some interesting properties of the resulting measures.

The Bank of England Survey of External Forecasters covers a sample of City firms, academic institutions and private consultancies, predominantly based in London. Their identities are not known to us, although an identification number allows us to track individual responses over time. In the beginning, in 1996, the quarterly survey asked for point and density forecasts of inflation on a fixed-target basis, namely the end-quarters of the current year and the next year. A fixed-horizon question, for forecasts two years ahead, was added in 1998, when GDP growth forecast questions were also added. May 2006 saw the two fixed-target questions replaced by fixed-horizon questions, one-year-ahead and three-years-ahead, so since that time we have quarterly series of one-, two- and three-years-ahead forecasts, for both variables, and these are the data studied in this paper. The dataset extends to November 2011 and thus comprises 23 surveys. The format of the questionnaire is illustrated by the recent inflation question shown in Figure 1. A summary of survey results is published in each quarterly Bank of England *Inflation Report*, most of whose content comprises the economic analysis and forecasts of the Monetary Policy Committee; we refer to the surveys by the month in which publication occurs (February, May, August, November), although the survey is undertaken in the last few days of the preceding month. Each published chart or table notes the number of responses on which it is based; this is typically in the low twenties, and varies by forecast horizon. This survey, like most others, thus faces the problems of item non-response as well as complete non-response.

The period since the questionnaire redesign in May 2006 includes the recent financial crisis, with more action in the data than in the period covered by our 2008 article, when the range of the quarterly observations of the then-target annual RPIX inflation rate was 1.9–3.2%. The Governor of the Bank of England famously described the earlier experience as ***n***on-***i***nflationary ***c***onsistently ***e***xpansionary, or ***'nice'***, (King, 2003), whereas recent experience has been rather different. To set the scene for the remainder of this paper, we show in Figure 2 the average central projections or point forecasts of the two variables over our sample period, as tabulated quarter-by-quarter in each *Inflation Report*, together with the latest data on the variable in question available to the forecasters at the time. Since such data relate to the previous quarter, we refer to the one-, two-, and three-years-ahead forecasts as having horizon h equal to 5, 9, and 13 quarters respectively. It is seen that, for both variables, the survey mean point forecasts show little reaction to current conditions. At longer horizons, the mean forecasts for CPI inflation and GDP growth scarcely deviate from the official inflation target rate of 2% and the trend growth rate of 2.5% respectively. However there is often considerable disagreement between individual point forecasts, as discussed below, which the mean forecasts conceal.

The remainder of the paper is organised as follows. Section 2 considers the statistical framework for the measurement and analysis of uncertainty, drawing some parallels with the literature on the elicitation of probability distributions. Section 3 compares several approaches to the construction of measures of uncertainty. For our preferred measure, we establish that there is substantial heterogeneity of forecast uncertainty, and significant persistence in individual relative uncertainty, using a statistical procedure based on rank orderings of individual uncertainty. In the appendix to the paper we replicate these findings for the ECB SPF one- and two-years-ahead density forecasts of inflation. Although relatively high or low uncertainty cannot be characterised as pessimism or optimism, these terms have been applied to similar findings of persistence in the relative level of individual point forecasts, using different statistical procedures, on different datasets. Accordingly we turn to the reported point forecasts in Section 4, and establish the same finding using our statistical procedure on the present dataset. Disagreement among individual point forecasts has been much discussed as a possible proxy for uncertainty in the absence of a direct measure of uncertainty, and we return to this question in Section 5, finding that the recent more turbulent period is rather more informative than the 'nice' period. It also illustrates the limitations of the survey average density forecast, regularly published in the *Inflation Report*, as an indicator of uncertainty. Section 6 concludes.

2. Measuring uncertainty: a framework for analysis

2.1. Elicitation and probability distributions

The literature on elicitation has some useful discussion of issues that are relevant to the collection and analysis of survey density forecasts. Elicitation is defined as the process of formulating a person's knowledge and beliefs about one or more uncertain quantities into a probability distribution for those quantities: see Garthwaite, Kadane and O'Hagan's (2005) review of the statistics literature, while the 2006 book by O'Hagan and seven colleagues provides more extensive coverage of the field.

A stylized representation of the elicitation process is as a dialogue between an *expert* or group of experts in the relevant subject matter, and a *facilitator*, who helps formulate expert knowledge in probabilistic form. In the case of a single unknown quantity, the process might begin by asking the experts to agree on its most likely value, the mode. They might then be asked to give probabilities for the variable lying in a small number of intervals, chosen to avoid asking them to assign very small probabilities, and to give good information around the mode. These are typically not presented as a histogram, but allow a histogram, usually with unequal bins, to be constructed. A continuous distribution, parametric or otherwise, is then fitted to the histogram, and its implied probabilities compared to those initially elicited. If these do not accurately represent the experts' beliefs, then iteration takes place until agreement is obtained.

Although the collection of histogram density forecasts from a panel of professional forecasters lacks the iterative nature of the elicitation process, the following observations are immediately relevant to both activities.

- It is clear that any probability distribution chosen to represent expert beliefs or to summarise a histogram density forecast expresses uncertainty about the variable in more detail than has been provided, and should not be interpreted as a perfect representation of expert or forecaster uncertainty.
- Reported probabilities are imprecise – there is uncertain uncertainty – but 'there cannot be a fully probabilistic solution to the problem of imprecision in probability assessments, as the notion of an imprecise probability itself is in violation of an axiom of subjective probability' (O'Hagan *et al*, 2006, p.160).
- The elicitation process and the reporting of a histogram density forecast are not processes of sampling from a population, and do not support the use of classical hypothesis tests of the goodness-of-fit of the chosen distribution.

These observations are elaborated in turn in the remainder of this section.

2.2. The use of probability distribution functions

The main use of fitted distributions in survey forecast research is to facilitate the estimation of measures of location and higher moments, as an alternative to the traditional method of estimating moments of distributions specified in histogram form. The traditional approach applies standard formulae for discrete distributions using representative values of the variable in each histogram bin. The representative values are usually taken to be the midpoints of the histogram intervals, adding an assumption about the highest and lowest intervals, which are open: these are usually treated as closed intervals of the same width or twice the width of the interior intervals. Thus the traditional approach treats the distribution as if all probability mass is located at the interval midpoints. However the underlying variable is continuous: the assumption that the probability distribution is uniform within each bin gives the same mean but increases the variance from the value given by the traditional approach by one-twelfth of the squared bin width. Typical macroeconomic density forecasts are unimodal, however, suggesting that more of the probability in each bin is located closer to the centre of the distribution. This motivates estimation of the variance by fitting normal distributions, which do not require the range of the variable to be restricted by an assumed closure of the open-ended highest and lowest intervals. This is applied to the individual and average US SPF density forecasts by Giordani and Soderlind (2003), whose Figure 3 shows that, for the survey average histograms, the fitted normal distributions imply smaller variances than those given by the traditional method.

Engelberg, Manski and Williams (2009) fit generalised beta distributions to the individual US SPF density forecast histograms. The generalised beta distribution is an ordinary beta distribution that is scaled to have support (L, U), where L and U are the bounds on the variable of interest. If the forecaster has placed non-zero probability in the upper and/or lower open intervals, then assumptions to close these intervals are required. With parameters $p>1$, $q>1$ the beta distribution matches the unimodal character of macroeconomic density forecasts, and in addition is very flexible (Johnson, Kotz and Balakrishnan, 1995, ch.25). If the implied point forecast is the focus of attention, for example, then the beta distribution allows different mean, median and mode, each of which can be defended as an appropriate point forecast. As is the case with the normal distribution, non-zero probabilities are needed in at least three bins to allow fitting to proceed; on occasions when a forecaster uses only two intervals, which are invariably adjacent, Engelberg *et al*. (2009) fit symmetric triangular distributions. Overall, they find that the reported point forecasts tend to be high percentiles of the probability distributions fitted to the density forecasts of GDP growth, and low percentiles in the case of inflation, indicating that forecasters tend to provide favourable point predictions relative to their probabilistic beliefs. This line of research is extended by Clements (2012), using the same distributional assumptions.

Two generalisations of the normal distribution are available to accommodate skewness or unbalanced risks in the forecast. One is the skew-normal distribution introduced by Azzalini (1985), fitted to the US SPF density forecasts by Garcia and Manzanares (2007). The other is the two-piece normal distribution originally due to Fechner (ed. Lipps, 1897), introduced to economic forecasting when the Bank of England's and Sveriges Riksbank's ‘fan chart’ forecasts of inflation appeared, in 1996. The Monetary Policy Committee (MPC) of the Bank of England was established a year later; it assumed responsibility for the inflation forecast and continued its presentation as a fan chart, published each quarter in the *Inflation Report*. Its production shares some features of the elicitation exercises discussed above.

The MPC presents its density forecasts of inflation graphically, as sets of forecast intervals covering 10, 20, 30,…, 90 per cent of the probability distributions, coloured red, and since the distributions become increasingly dispersed and the intervals ‘fan out’ as the forecast horizon increases, the chart soon became known as the ‘fan chart’. An example from the February 2010 *Inflation Report* is shown in Figure 3. Selected cross-sections of the fan chart are also published, such as the two-years-ahead forecast from Figure 3 shown in Figure 4. The histogram is drawn using shortest or equal height intervals, so the darkest red band contains the most likely outcome, the mode of the distribution. The balance of risks is skewed to the upside in this example, with the 10 per cent probability in each pair of lighter red bands divided 4:6, thus the area of the histogram bars is constant at 0.04 (lower) or 0.06 (upper) and the histogram intervals are all different. It is known that the Bank staff (the ‘facilitators’) use the two-piece normal distribution to calculate the histogram intervals, but the histogram is all that is agreed by the MPC (the ‘experts’). The recent comment by the Governor, that ‘the distribution of … tail events is not explicitly specified, as to do so would require a spurious degree of precision on the part of the MPC’ (King, 2010, p.14), is then in complete accordance with our first bullet point above. Recognition of this point possibly lay behind the decision in August 2006 to show the cross-sections as histograms, as in Figure 4: before this the cross-sections were shown as continuous densities.

2.3. Uncertain uncertainty

The elicitation literature discusses two main reasons for exercising caution in interpreting any elicited probability distribution as a perfect representation of an expert’s uncertainty. One is that the relatively few probability assessments provided by the expert are not sufficient to specify a unique probability distribution; there are many different distributions that would fit them exactly. It is possible, however, that the quantity of interest derived from the distribution, which in our present case is a measure of its dispersion, is robust to changes in the chosen distribution, and this is explored in Section 3.

The second reason for caution is that it is difficult for experts to give precise numerical values for their probabilities. Oakley and O'Hagan (2007) allow for 'noise' in an expert's reported probabilities by postulating that each reported probability comprises the expert's 'true' probability plus an additive error that represents the imprecision in the reported probability. The error is then assumed to be normally distributed with mean zero and variance that depends on the reported probability, being smaller when this is closer to 0 or 1. In a Bayesian framework, Oakley and O'Hagan show how adding imprecision allows the facilitator's posterior distribution to assume a smooth, unimodal form in an example where the expert's elicited probabilities by themselves do not admit this possibility. In a classical framework, goodness-of-fit tests of the fitted distribution require an underlying model of the source and nature of random variation, on which such tests would be conditional: Oakley and O'Hagan's model is a possible example. However it is immediately clear on looking at forecast survey data that forecasters' uncertainty about their subjective probabilities does not match such a model, but is instead demonstrated by their widespread yet varying use of round numbers.

In statistical reporting, the level of rounding implicitly conveys information about the quality or accuracy of the data, and so aids interpretation of the data. Similarly, individual forecasters' level of rounding of their reported probabilities conveys information about the subjective uncertainty inherent in their probability assessments, which varies across individual forecasters. The overall extent of rounding is indicated in Table 1, which shows the proportion of non-zero percentage probabilities of different numerical characteristics observed in the complete dataset – all respondents, all histogram bins, both variables, and three forecast horizons, almost fifteen thousand numbers in total. It is seen that almost two-thirds of all bin probabilities are even or odd multiples of 5. Multiples of 10 are used in every

Table 1 The numerical character of reported probabilities

Reported percent probability	Percentage of cases
Multiple of 10	35.2
Otherwise multiple of 5	30.5
Other integer	30.7
Non-integer	3.6

bin of the histogram in approximately ten per cent of the individual forecasts. More commonly, round numbers are used in the centre of the distribution and smaller numbers in the tails, which also include most of the non-integer values recorded in Table 1. There is thus considerable variation in the treatment of uncertain probability assessments, both in the pattern of rounding across the distribution, and in the extent of rounding by individual respondents. Nevertheless the extent and pattern of rounding by individual survey respondents has strong similarities across the three forecast horizons, the two variables, and over time, indicating a feature of forecast methods or forecaster characteristics that is worthy of further investigation.

In contrast to percentage probabilities stated as round numbers, the reported probabilities of a small number of survey respondents appear to be calculated from a known probability distribution. The normal distribution, with possible extensions to asymmetry noted above, is a convenient and popular choice among forecasters who publish density forecasts. The distribution is typically centred on the associated point forecast, with variance calibrated to recent past point forecast errors, possibly with judgmental adjustment. Thus the separate roles of expert and facilitator discussed above are combined in a single forecaster or forecast team. In such cases the estimation of forecast uncertainty, to which we turn next, simply amounts to the recovery of the parameter values used by the forecaster.

3. The properties of individual uncertainty measures

3.1. Choosing a measure

We first report the results of a comparative study of variance estimates based on the normal and beta distributions. In each case the parameters are estimated by fitting the cumulative distribution function to the cumulated density forecast histogram by least squares, and we begin by discussing the results for the inflation forecasts.

To fit either distribution it is required that the histogram has at least three bins to which non-zero probabilities have been assigned by the forecaster. Of the available 1412 individual forecasts of inflation this requirement excludes 31 cases, in which only two bins are used, and we adopt the triangular distribution of Engelberg *et al*. (2009) in these cases. In the great majority of the remaining cases, including most of the three-bin cases, the histograms plotted assuming a constant bin-width are unimodal on an interior bin, and the two distributions have very similar goodness-of-fit and give very similar estimates of the variance, which are smaller than the estimate given by the traditional calculation of moments. The exceptions to this generalisation allow some discrimination between the distributions, however, which is partly related to shortcomings in the histogram designs.

The seven-bin inflation histogram design shown in Figure 1 has been in use since the February 2009 survey. Previously, six bins were specified, with four half-percentage-point bins covering the range 1–3%, and open bins above and below this range. During 2008 there was much public discussion of the possibility of below-target inflation, if not deflation, and in the survey report published in the November 2008 *Inflation Report* the average percentage probability in the lower open bin was, for the first time, in double digits, this occurring at all three forecast horizons. We believe that this prompted the division of the lower bin by the survey managers, to give the new format shown in Figure 1. However inflation prospects changed in 2010, and we began to see double-digit survey average percentage probabilities in the upper open bin, approaching 20% in the one-year-ahead forecasts in August and November 2010. But no change to the histogram design resulted. The consequence of these two episodes is that there are several individual forecast histograms which, when plotted with closed bins replacing the open bins, have a U-shaped or J-shaped appearance. The beta distribution can match these shapes, with parameters less than 1, but this conflicts with our strong prior that the underlying forecast density has a single interior mode. Changing the support of the beta distribution can moderate some of these cases, but such changes to the range of the variable are arbitrary. On the other hand the normal distribution imposes its unimodal shape in a manner determined by the available interior observations, and this is our preferred solution.

As an illustration, we choose our most extreme three-bin case, which is an individual one-year-ahead forecast in the August 2010 survey, with percentage probabilities of inflation in the ranges 2–2.5%, 2.5–3%, and >3% of 5, 5, and 90 respectively. How to choose a sensible upper bound that would allow a beta distribution with an interior mode to be fitted is an open problem. However, the normal distribution whose 5^{th} and 10^{th} percentiles are respectively 2.5 and 3.0 is an acceptable representation, with mean equal to 4.77 and standard deviation equal to 1.38, which is our preferred measure of uncertainty. We note that this respondent's reported central projection of CPI inflation was 4.0%, the highest among all survey respondents, and that the inflation outcome in 2011Q3 was 4.7%.

Turning to the GDP growth forecasts, we find further limitations in the histogram design. Over the period to August 2008 four bins were specified (<1%, 1–2%, 2–3%, >3%), but as recession fears increased the survey average probability observed in the lower bin increased, reaching 38% for the one-year-ahead forecasts in August 2008, so this bin was divided for the November 2008 survey, and again in February 2009, since when six bins have been specified (<−1%, −1–0%, 0–1%, 1–2%, 2–3%, >3%). At the upper end of the range, with a long-run trend growth rate above 2% per annum, it is no surprise that the survey average probability of growth exceeding 3% in the three-years-ahead forecasts is often in excess of 25%, but no comparable changes to the configuration of the bins have been made. The consequence is that we have more individual forecasts with J-shaped histograms in the growth forecasts, peaking in one or other open bin, than in the inflation forecasts. Moreover the use of wider interior intervals results in more cases in which there are only two bins to which non-zero probabilities have been assigned by the forecaster (59 out of 1404 available forecasts), also seven cases with 100% probability assigned to a single interior bin. For the two-bin cases we continue with the variance estimate from the triangular distribution of Engelberg *et al*. (2009), as above, and for the one-bin cases we set the variance equal to $\frac{1}{24}$, assuming a symmetric triangular distribution with unit support. The general results of a comparison between the beta and normal distributions for the GDP growth forecasts are otherwise very much in line with those obtained for the inflation forecasts, described above, but there are more exceptional cases for which the beta distribution is inappropriate. Hence our preferred uncertainty measure for both variables is the standard deviation of the fitted normal distribution, except in one- and two-bin cases, as noted above.

The resulting estimates of individual standard deviations are shown in Figure 5, whose six panels refer to the two variables and three forecast horizons. As a point of reference the solid line in each panel shows the median individual standard deviation, around which we observe substantial dispersion. Note that the scale of the inflation panels is different from that of the GDP growth panels: the latter variable is generally considered to be more difficult to forecast, not least due to the problems caused by data revisions, whereas the CPI is never revised after first publication. The general level of uncertainty is lower at shorter forecast horizons, as expected. The median measure has a local peak in February 2009, which in some cases appears to signal a shift in level. The general spread of uncertainty measures also appears to have increased from that time. Figure 2 shows that February 2009 did not mark a turning point in either of the underlying series. Rather, these forecast measures reflect increased uncertainty about economic prospects as the global financial crisis spread, and central banks made unprecedented policy moves, with the MPC cutting UK bank rate by 3 percentage points between the November 2008 and February 2009 forecast surveys. Similarly substantial heterogeneity of forecast uncertainty, varying over time, is demonstrated over 35 years of the US SPF by Lahiri and Liu (2006, Fig.1).

3.2. The persistence of individual relative uncertainty

On moving to more systematic study of individual characteristics we immediately face the missing data problem. The Bank of England survey, like other forecast surveys and individual panel studies more generally, has experienced exit and entry of participants and sporadic non-response, both to the complete questionnaire and to items within it: the longer-horizon forecasts are more often missing than the one-year-ahead forecasts. To avoid the complications caused by long gaps in the data we follow common practice in survey research and conduct our analysis of individual forecasters on a subsample of 'regular' respondents. In this paper the subsample comprises the 17 respondents whose item response rate over the 23 surveys exceeds two-thirds. The overall subsample item response rate is 87%; over six questions and 23 surveys the number of available responses ranges between 13 and 17.

To study possible persistence in individual forecasters' relative uncertainty, we identify the regular respondents in Figure 5, and for each column in each panel of the figure, rank them from the highest to the lowest uncertainty. In each panel we next calculate each forecaster's average rank over the (23 or fewer) surveys in which they appear. To illustrate possible persistence in relative positions within the panels of Figure 5, we then delete all but the five highest-ranked and the five lowest-ranked individuals, and show their uncertainty measures in Figure 6, respectively in blue and red. Note that missing observations imply that there are typically less than ten points in each column; also these are not the same ten individuals across the six panels, a question we return to below. To centre the spread, we retain the plots of the median individual uncertainty from Figure 5.

There is a very clear indication of persistence in relative forecast uncertainty in Figure 6, with blue dots tending to stay high and red dots tending to stay low in all six panels. Some of the obvious outliers in Figure 5 do not appear in Figure 6: most of these deletions are the uncertainty measures of one non-regular respondent; the others are due to a regular respondent whose average rank is not in the highest or lowest five. An eye-catching case is the single respondent whose uncertainty is the highest shown, at a constant level, over the first ten surveys in our sample, in the inflation, $h=9$ and 13, and GDP growth, $h=13$ panels of Figure 6. In the next survey, November 2008, there was no response, then responses resumed, but no longer occupying the top position, and with a different pattern of rounding, suggesting that there had been a change of forecast personnel, models or methods (or all three) in this institution.

A statistical measure of the similarity over time of the rankings in each panel of these figures, and hence of the persistence of individual relative uncertainty, is provided by the Kendall coefficient of concordance (Kendall and Gibbons, 1990, ch.6). Usually denoted W, this is defined as the ratio of the sum of squared mean deviations of the observed average ranks to its maximum possible value, thus $0 \le W \le 1$. With 17 regular respondents and no missing data, perfect agreement of the rankings across all 23 surveys would give average ranks 1, 2,…,17 in some order, with sum of squared mean deviations equal to 408, which is the maximum possible value in this case. At the other extreme the average ranks all tend to equal 9 when rankings of individuals are purely random over time.

Whenever observations are missing, however, the maximum possible rank is less than 17. For each survey we rank the forecasts ignoring non-respondents, and individuals' average ranks are calculated over the occasions on which they responded. We calculate a revised maximum possible sum of squared mean deviations of average ranks conditional on the observed pattern of missing data in each case, and with the observed average ranks we obtain the results shown in Table 2. To aid interpretation of these coefficients, we note that, under a null hypothesis of random rankings over time, with r rankings of n individuals and no missing data, $r(n-1)W$ is approximately distributed as chi-squared with $n-1$ degrees of freedom, hence the 95th and 99th percentiles of W in these circumstances, with $r=23$ and $n=17$, are 0.07 and 0.09 respectively. The coefficients in Table 2 thus indicate considerable stability over time in the relative level of individual forecasters' uncertainty, to a similar

Table 2 Measures of agreement over time between forecasters' rankings with respect to their uncertainty measures: Kendall coefficients of concordance

	$h=5$	$h=9$	$h=13$
CPI inflation	0.40	0.50	0.47
GDP growth	0.44	0.40	0.47

extent for both variables and all three forecast horizons. In order to pool these cases, we calculate the concordance between the six rankings given by the time-averaged scores in each case. The Kendall coefficient is 0.92; under the above null its 99th percentile for 6 rankings of 17 individuals is 0.33. The rankings of individual forecasters by their uncertainty levels are almost identical across the two variables and three forecast horizons we consider.

This strong evidence of persistence in individual forecasters' relative levels of uncertainty, as expressed in their subjective probabilities, is a new finding. We are not aware of models of forecaster behaviour that incorporate this feature. Research to date has focused on differences in point forecasts, and in Section 4 we use the statistical procedures employed above to replicate some existing findings of persistence in the relative levels of individual point forecasts. Before that we report a preliminary exploration of these differences in individual forecasters' ex ante uncertainty.

3.3. Ex ante and ex post measures of inflation uncertainty

Published density forecasts of inflation in the UK date from February 1996, which marks the first appearances of the Bank of England's fan chart, noted above, and the National Institute of Economic and Social Research's histograms, as tables. NIESR also published a density forecast of real GDP growth at this time, whereas a growth fan chart did not appear in the Bank's *Inflation Report* until November 1997. In both institutions the variance of the density forecast was calibrated with reference to past point forecast errors, with judgmental adjustment. This practice has continued, although the reduction in inflation variance during the 1990s was not recognised quickly enough (Wallis, 2004; Mitchell, 2005; Bank of

England, 2005). This prompts us to investigate whether forecast survey respondents behave in a similar way or, more specifically, whether our ex ante density forecast uncertainty measure is related, at the individual level, to an ex post uncertainty measure based on past point forecast errors.

We measure individual forecasters' ex post uncertainty by their point forecast RMSE over the preceding four quarters. Since the quarterly series of one- and three-years-ahead forecasts began only in May 2006, the need to wait for four forecast outcomes in order to obtain an RMSE observation makes the available time-series samples unacceptably small. However the two-years-ahead forecast questions were part of the questionnaire before May 2006, so for these forecasts we have sufficient past data to utilise the full sample, with T=23, subject to individual forecasters' missing observations. We consider only the inflation forecasts, in view of the ambiguities that surround GDP growth forecasts and their evaluation as a result of the large data revisions to which this variable is subject.

Joint estimation of a regression of density forecast standard deviation on point forecast RMSE for 17 regular respondents, allowing intercept and slope to vary across individuals, yields strong rejections of the equality of intercepts and of the equality of slope coefficients. The first of these is another reflection of the persistence in relative uncertainty already noted. With respect to the individual slope coefficients, eight are positive and significantly different from zero, and eight are positive but not significantly different from zero. In the remaining case there is a significant negative coefficient, which is clearly due to the presence of two distinct subsamples in the data: this is the case of an apparent institutional change noted above. Once again we have evidence of different forecasters following different practices: some do appear to calibrate their density forecasts with reference to past point forecast errors; some do not, at least with the measure we have chosen. If we neglect this heterogeneity and impose equality of the individual coefficients on point forecast RMSE, then the resulting common coefficient is significantly different from zero, in contrast to the results of Lahiri and Liu (2006) with US SPF data, also those of Rich and Tracy (2003) that they cite. However Rich and Tracy were working with survey average data, while Lahiri and Liu used the level and absolute value of the immediate past forecast error rather than a forecast RMSE measure.

4. Persistence in the relative level of individual point forecasts

We first note two precursory studies based on data from the *Consensus Economics* service, which is a monthly survey of private sector forecasting bodies in a number of countries. Forecasts are collected for the current year and the following year, so forecasters eventually supply a series of 24 forecasts for each target year. Several researchers have made use of the individual point forecasts of GDP growth and CPI inflation. Batchelor (2007) finds persistent individual biases towards optimism or pessimism in GDP growth forecasts for the G7 countries at all horizons for the target years 1991-2004; for the inflation forecasts there is less consistency across countries and forecast horizons. Patton and Timmermann (2010) study the persistent behaviour of US forecasters by classifying their point forecasts in each year into three groups – high, medium, low – and studying the associated Markov transition matrix. This is done separately for 'short-horizon' (current year) and 'long-horizon' (following year) forecasts of GDP growth and CPI inflation: there is a significant tendency for forecasters to stay in the same group from one year to the next.

The individual point forecasts from the Bank of England's survey are shown in the six panels of Figure 7, together with the survey mean point forecasts shown in Figure 2. The overall dispersion of individual forecasts around the mean is relatively small until early 2009, when there is considerable disagreement about the consequences of the crisis and the policy actions taken in response to it. To study possible persistence in the relative positions of individual forecasters' point forecasts, we repeat the transition between Figures 5 and 6 seen above, again working with the 17 regular respondents. Thus we first delete non-regular respondents from Figure 7, and for each time-period in each panel rank the regular respondents from the highest to the lowest point forecast. Next, in each panel we calculate each forecaster's average rank over the periods for which they supplied a forecast. Finally we delete the forecasts of all respondents except those with the five highest and five lowest average ranks, plotted in blue and red respectively, to obtain Figure 8.

As in Figure 6 there are clear indications of persistence, which in Figure 8 relate to the relative levels of point forecasts. Again there is an overall tendency for blue dots to stay high and red dots to stay low across each panel of Figure 8, which is more pronounced in the GDP growth forecasts, and less so in the inflation forecasts, in particular when the dispersion is relatively small. This visual assessment is confirmed by the Kendall coefficients of

Table 3 Measures of agreement over time between forecasters' rankings with respect to their point forecasts: Kendall coefficients of concordance

	$h=5$	$h=9$	$h=13$
CPI inflation	0.27	0.27	0.15
GDP growth	0.42	0.46	0.53

concordance presented in Table 3. The persistence in the relative level of point forecasts of GDP growth is similar to that observed in the relative level of forecast uncertainty; for inflation, the persistence in relative point forecasts is less than that in the uncertainty measures. In comparing Figure 8 and Table 3 we note that Table 3 is based on the rankings of point forecasts and not on the forecast inflation or growth rates, hence in the inflation $h=13$ panel of Figure 8, for example, the highest and lowest observations in May–November 2009 receive the same weight in the concordance calculations as the highest and lowest observations in August 2007–May 2008, although the separation of these earlier observations is much smaller. The frequent interchange of position within this reduced spread contributes to the relatively small value of the coefficient in this case.

Again we pool the six cases represented in the panels of Figures 7 and 8 and the cells of Table 3, and calculate the concordance between the six rankings implied by the time-averaged scores for each variable/forecast horizon combination, inverting the GDP rankings: the Kendall coefficient is 0.46. This is one-half of the value obtained in the previous section, comparing the rankings of forecasters by their uncertainty levels across the two variables and three forecast horizons, nevertheless it is still well clear of the 99th percentile under the null, of 0.33, indicating strong similarity of these point forecast average rankings. Neither of the articles cited above considers a bivariate notion of optimism (forecasts of low inflation ***and*** high growth) and pessimism (vice versa) in analysing the *Consensus Economics* data. This result takes us a step further, showing that, in the Bank of England survey, forecasters with relatively low point forecasts of inflation tend to have relatively high forecasts of GDP growth, and vice versa, persistently so over this period.

5. Uncertainty and disagreement revisited

Disagreement among point forecasts is often used as an indicator of uncertainty in the absence of a direct measure. The availability of measures of uncertainty based on density forecasts allows the utility of such proxy variables to be assessed, and a research literature has developed, mostly based on US SPF data, originating with the seminal article by Zarnowitz and Lambros (1987). We studied this question using the Bank of England survey, 1996-2005, in our 2008 article, and we return to it now with the present dataset, which covers a less quiescent period, and with a better measure of uncertainty, as developed above.

A framework for analysis is provided by an expression for the variance of the survey average density forecast (Wallis, 2005). Denoting n individual density forecasts of a random variable Y at some future time as $f_i(y)$, with mean μ_i and variance σ_i^2, $i = 1,...,n$, the survey average density forecast is

$$f_A(y) = \frac{1}{n}\sum_{i=1}^{n} f_i(y).$$

Its mean and variance are

$$\mu_A = \frac{1}{n}\sum_{i=1}^{n}\mu_i, \qquad \sigma_A^2 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 + \frac{1}{n}\sum_{i=1}^{n}(\mu_i - \mu_A)^2.$$

These expressions hold irrespective of the forms of $f_i(y)$, which might include histograms, as in the surveys. The last equation says that the variance of the survey average density forecast is equal to the average individual uncertainty (variance) plus a measure of the dispersion of, or disagreement between, the individual density forecast means. If these last means are being used as point forecasts, then we have an exact relation involving the disagreement between them. The practical situation under consideration, however, is one in which only point forecasts, from diverse sources, are available, and using a measure of their disagreement introduces a discrepancy into the equation. A second practical feature is that it is usually preferred to report standard deviations, not variances, in order to show measures whose units coincide with the units of the variable under consideration, whereas the equation holds for variances, and not for standard deviations. It nevertheless provides a useful conceptual framework for the problem, in particular making clear why the dispersion of the survey average density forecast is not an appropriate indicator of aggregate uncertainty.

Our practical measures corresponding to the three terms in the equation are then
(i) the standard deviation of the survey average density forecast, estimated via a fitted normal distribution;
(ii) the square root of the average of the individual variances estimated as above, termed the root mean subjective variance (*RMSV*) in our 2008 article, following Batchelor and Dua (1996);
(iii) a robust quasi-standard deviation (*qsd*) measure of disagreement, also used in our 2008 article, following Giordani and Soderlind (2003), given as one-half of the difference between the 16th and 84th percentiles of the sample of point forecasts, appropriately interpolated. (For a normal distribution, this interval is equal to the mean ± 1 standard deviation.)
These three measures are plotted in the six panels of Figure 9; note that, as in several preceding figures, different scales are used for the inflation and GDP growth panels.

In all six panels of Figure 9 the experience over the first two years shown is very similar to that seen over the earlier years of this decade in our previous article, with low disagreement and relatively little movement in the series, hence little possibility of interesting co-movements. The picture then changes dramatically with the onset of the crisis, as suggested in some of the preceding figures. There are rapid increases in all three series in all six panels, most prominently so in the disagreement measure, as anticipated in Figure 7. These pronounced movements in common result in the high correlations shown in Table 4, and the conclusion that, over this period, changes in disagreement are associated with changes in uncertainty. The changes in uncertainty are proportionately smaller in all cases; in the four cases with prominent spikes in disagreement they contribute less to changes in the variance of the survey average density forecast than the changes in the disagreement measure.

Table 4 Correlation coefficients between uncertainty (*RMSV*) and disagreement (*qsd*)

	$h=5$	$h=9$	$h=13$
CPI inflation	0.85	0.81	0.65
GDP growth	0.59	0.67	0.81

The switch from a negative to a positive answer to the question, is disagreement a useful proxy for uncertainty, between our earlier article and the present work is a mirror image of research findings on the US SPF data. Zarnowitz and Lambros originally found 'some direct empirical support … that greater interpersonal differentiation of expectations is a symptom of greater uncertainty' (1987, p.607), using data from the survey's start, in late 1968, to 1981. In contrast, Lahiri and Liu (2006) and Rich and Tracy (2010) give negative answers, both articles being based on a much longer sample, from 1968 to the early 2000s. The period from 1968 to 1981 is dominated by the Great Inflation, while the longer period adds in the Great Moderation; a significant reduction in the volatility of US inflation and output in the early 1980s has been widely documented. Thus the joint results from the US and UK surveys suggest the encompassing conclusion that disagreement is a useful proxy for uncertainty when it exhibits large fluctuations, but low-level high-frequency variations are not sufficiently correlated. We await updating of the US studies to the recent crisis period.

6. Conclusion

In this paper we consider several statistical issues that arise in the construction and interpretation of measures of forecast uncertainty from individual histogram density forecasts obtained by surveying forecasters. We find substantial heterogeneity in forecasters' uncertainty about future outcomes, as expressed in their subjective probabilities, and strong persistence in the relative level of individual forecasters' uncertainty. This is a new finding, demonstrating individual characteristics of forecaster responses that merit deeper investigation. Using the same statistical procedures we also find persistence, at a lower level, in individual forecasters' relative point forecasts, reflecting their relative optimism or pessimism about future prospects for inflation and GDP growth. This is not a new finding for these variables taken separately, but we also establish that a bivariate relation exists, jointly defining a persistently optimistic forecast as one of relatively low inflation and high growth, and pessimism vice versa.

Our experience in conducting this research also leads to suggestions for improving the reporting of survey results. Since the available samples of these demonstrably heterogeneous forecasters are not large, and vary over time, comparison of summary results between the current and preceding surveys, which is a common practice, can be strongly affected by missing observations in one or other survey. Our recommendation is that such comparisons be based only on the individual respondents who are present in both surveys. Secondly, although information about disagreement is often supplied, as a histogram of point forecasts, for example, (also an example of a ‘this quarter/last quarter’ comparison just mentioned), little is typically reported about uncertainty. A table of survey average probabilities is standard, but its only use might be to describe the survey average probability that future inflation or growth will lie to one side or the other of a threshold of interest. However a measure of average individual uncertainty could be derived, without replicating all our calculations described above, as the difference between a histogram-based variance of the survey average density forecast and the variance of the histogram of point forecasts (disagreement): the variance equation in Section 5 above then yields the implied mean subjective variance.

Appendix. Persistent uncertainty in the ECB Survey of Professional Forecasters

This appendix describes our replication of the analysis presented in Section 3 above on the forecasts of euro-area inflation collected by the European Central Bank. The ECB’s Survey of Professional Forecasters began in 1999 at the same time as the establishment of the euro area and its central bank. Aggregate macroeconomic variables had not previously been widely available on an area-wide basis, so we allow a two-year training period for forecasters and data compilers, and work with a series of 44 quarterly surveys from 2001Q1 to 2011Q4. Over this period the number of individual responses varies between 34 and 56, and we analyse their forecasts of HICP inflation one- and two-years-ahead. As in the main text, we set the scene by showing in Figure A1 the survey mean point forecasts together with the latest inflation data available to the forecasters: as in Figure 2, the prominent spike in inflation was correctly believed to be temporary.

For the density forecasts, the histogram design specifies closed half-percentage-point bins over a wider range of inflation than that covered in the Bank of England’s questionnaire. Initially this range was 0−3.4% (7 bins), with open bins above and below. Subsequently the open bins have been divided, and currently there are 10 closed bins covering the range –1.0−3.9%, with open bins above (>4.0%) and below (<−1.0%). Unusually, the bins are specified discontinuously, as … 1.0–1.4%, 1.5−1.9%, 2.0−2.4%, … , and different respondents treat the gaps differently. We assume that the underlying distribution is continuous, and in fitting a distribution in order to obtain an estimate of an individual respondent's variance it makes no difference whether we specify that the steps in the cumulative distribution occur at values ending in .0 and .5 or at values ending in .4 and .9. The general level of uncertainty about inflation prospects is lower in the euro area, which results in the use of fewer bins than in the Bank of England survey. Of the 3921 individual forecasts, there are 450 cases in which only two bins have non-zero probabilities, and 13 cases with 100% probability assigned to a single bin. The specification of closed bins over a relatively wide range then results in very little use of the open bins by survey respondents, and in particular there are no U-shaped or J-shaped histograms.

Turning to estimation of standard deviations, in the one- and two-bin cases we continue with the assumption of a triangular distribution as above; in the one-bin case we assume that its support is 0.5 so the standard deviation is $\sqrt{1/96} = 0.102$. For the remaining cases estimates based on fitted beta and normal distributions are very close to one another, such that the eye can scarcely distinguish between their time-series plots, and we choose to continue with the estimates from the normal distribution. The estimates for all individual observations are shown in the upper panels of Figure A2. The solid line is the median individual standard deviation, which is generally lower for the shorter horizon forecasts, with both approximately constant until mid-2008, then increasing to a peak in mid-to-late 2009. There is substantial dispersion in general, also increasing around this time.

To replicate the statistical analysis in Section 3.2 we likewise proceed with a subsample of regular respondents. Here these are the 26 respondents whose item response rate for these two forecasts and 44 surveys exceeds 70%. The overall subsample item response rate is 87%; the number of available responses on any single forecast never falls below 18. To study possible persistence in individual relative uncertainty we repeat the exercise of Section 3.2, identifying the regular respondents in the upper panels of Figure A2, ranking them from the highest to the lowest uncertainty quarter-by-quarter across each panel, and calculating the resulting average ranks in each panel. We select the eight highest-ranked and the eight lowest-ranked individuals and retain their uncertainty measures in the lower panels of Figure A2, respectively in blue and red. Figure A2 gives a stronger indication of persistence in relative forecast uncertainty than that seen in Figure 6, for the Bank of England survey. This visual impression is confirmed by the Kendall coefficients of concordance between the rankings over time within each panel, which for the one- and two-years-ahead forecasts are 0.64 and 0.63 respectively. These exceed the corresponding coefficients presented in Table 2 while, with increases in the numbers of respondents and time periods, the critical values under the null hypothesis are reduced. The correlation between the overall rankings for each forecast horizon is 0.97. This again exceeds the corresponding correlation in the Bank of England survey which, for the rankings in the inflation $h = 5$ and 9 panels of Figure 6, is 0.84. In the ECB data the rankings of individual forecasters by their uncertainty levels are almost identical across the two forecast horizons.

Finally, we turn to the question posed in Section 3.3, whether there is evidence that forecasters calibrate their density forecasts with reference to recent past point forecast errors. The test is the same, namely time-series regression of density forecast standard deviation on point forecast RMSE over the preceding four quarters. We record point forecasts from the beginning of our sample period, 2001Q1, hence the first one-year-ahead point forecast errors are observable a year later, and our series of point forecast RMSE based on four forecast errors, and hence our regression sample period, begins in 2003Q1; for the two-years-ahead forecasts the regression sample period begins a year later. In each case joint estimation of the individual regressions for 26 regular respondents yields strong rejections of the equality of intercepts and the equality of slope coefficients . The first again reflects the persistence in relative uncertainty levels. With respect to the individual slope coefficients in the one-year (two-years) ahead forecast regressions, 15 (18) coefficients of past RMSE are positive and significantly different from zero, nine (six) are positive and not significantly different from zero, and two (two) are negative and insignificantly different from zero. As above, there is evidence of heterogeneity in forecasters' practice. If this is neglected, and the equations estimated subject to equality of the individual coefficients of point forecast RMSE, then the resulting common coefficient is again significantly different from zero, as in the main text. Indeed, all of the results of Section 3 have been successfully replicated with the ECB SPF data.

References

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171-178.

Bank of England (2005). The MPC's forecasting record. *Inflation Report*, August 2005, pp.40-41.

Batchelor, R.A. (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting*, 23, 189-203.

Batchelor, R.A. and Dua, P. (1996). Empirical measures of inflation uncertainty: a cautionary note. *Applied Economics*, 28, 333-341.

Boero, G., Smith, J. and Wallis, K.F. (2008). Uncertainty and disagreement in economic prediction: the Bank of England Survey of External Forecasters. *Economic Journal*, 118, 1107-1127.

Clements, M.P. (2012). Probability distributions or point predictions? Survey forecasts of US output growth and inflation. Warwick Economic Research Papers No.976.

Engelberg, J., Manski, C.F. and Williams, J. (2009). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business and Economic Statistics*, 27, 30-41.

Fechner, G.T. (ed. Lipps, G.F.) (1897). *Kollektivmasslehre*. Leipzig: Engelmann.

Garcia, J.A. and Manzanares, A. (2007). What can probability forecasts tell us about inflation risks? Working Paper No.825, European Central Bank, Frankfurt.

Garthwaite, P.H., Kadane, J.B. and O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100, 680-700.

Giordani, P. and Soderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, 47, 1037-1059.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, vol.2 (2nd ed.). New York: Wiley.

Kendall, M.G. and Gibbons, J.D. (1990). *Rank Correlation Methods* (5th ed.). London: Edward Arnold.

King, M.A. (2003). Speech to the East Midlands Development Agency, 14 October 2003. http://www.bankofengland.co.uk/publications/Documents/speeches/2003/speech204.pdf

King, M.A. (2010). Uncertainty in macroeconomic policy making: art or science? Royal Society Conference, 22 March 2010. http://www.bankofengland.co.uk/publications/Documents/speeches/2010/speech432.pdf

Lahiri, K. and Liu, F. (2006). Modelling multi-period inflation uncertainty using a panel of density forecasts. *Journal of Applied Econometrics*, 21, 1199-1219.

Mitchell, J. (2005). The National Institute density forecasts of inflation. *National Institute Economic Review*, No.193, 60-69.

Oakley, J.E. and O'Hagan, A. (2007). Uncertainty in prior elicitations: a nonparametric approach. *Biometrika*, 94, 427-441.

O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E. and Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*. Chichester: Wiley.

Patton, A.J. and Timmermann, A. (2010). Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, 57, 803-820.

Rich, R. and Tracy, J. (2003). Modeling uncertainty: predictive accuracy as a proxy for predictive confidence. Staff Report No. 161, Federal Reserve Bank of New York.

Rich, R. and Tracy, J. (2010). The relationships among expected inflation, disagreement, and uncertainty: evidence from matched point and density forecasts. *Review of Economics and Statistics*, 92, 200-207.

Soderlind, P. (2011). Inflation risk premia and survey evidence on macroeconomic uncertainty. *International Journal of Central Banking*, 7, 113-133.

Wallis, K.F. (2004). An assessment of Bank of England and National Institute inflation forecast uncertainties. *National Institute Economic Review*, No.189, 64-71.

Wallis, K.F. (2005). Combining density and interval forecasts: a modest proposal. *Oxford Bulletin of Economics and Statistics*, 67, 983-994.

Zarnowitz, V. and Lambros, L.A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, 95, 591-621.

Figure 1. Bank of England questionnaire, November 2010 survey, inflation question

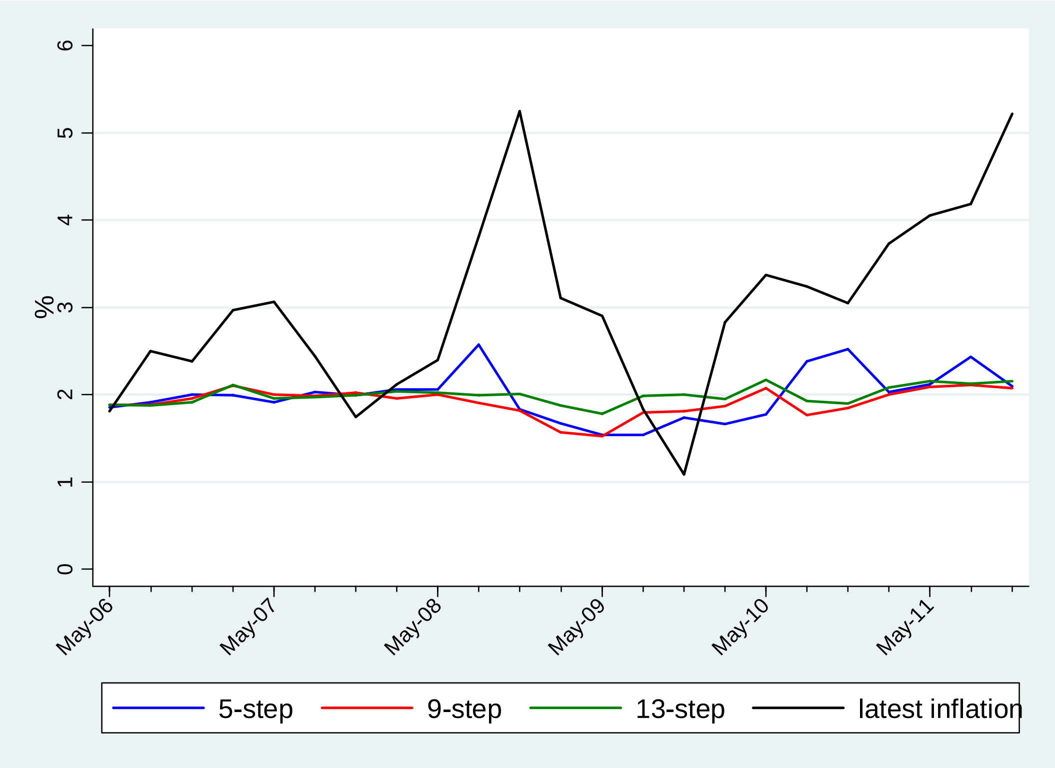
PROBABILITY DISTRIBUTION OF 12-MONTH CPI INFLATION OVER THE MEDIUM TERM

Please indicate the percentage probabilities you would attach to the various possible outcomes in 2011 Q4, 2012 Q4 and 2013 Q4. The probabilities of these alternative forecasts should of course add up to 100, as indicated.

PROBABILITY OF 12-MONTH CPI INFLATION FALLING IN THE FOLLOWING RANGES			
	2011 Q4	2012 Q4	2013 Q4
<0%			
0.0% to 1.0%			
1.0% to 1.5%			
1.5% to 2.0%			
2.0% to 2.5%			
2.5% to 3.0%			
> 3.0%			
TOTAL	100	100	100

CENTRAL PROJECTION FOR 12-MONTH CPI INFLATION		
2011 Q4	2012 Q4	2013 Q4

Figure 2. Mean point forecasts, and latest (monthly) data available to forecasters
Upper panel: CPI inflation; lower panel: GDP growth



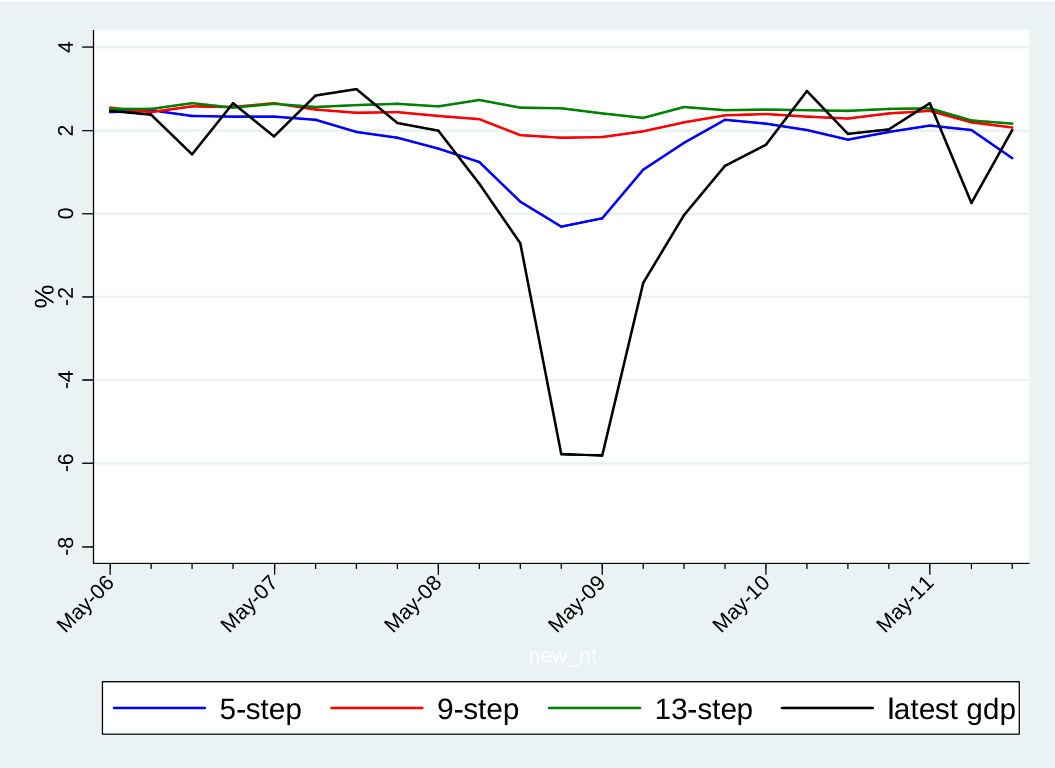


Figure 3. Bank of England MPC fan chart forecast of inflation, February 2010

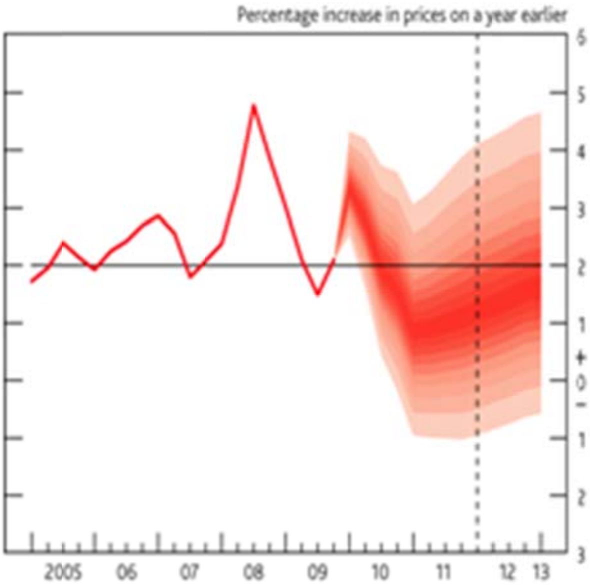


Figure 4. Two-years-ahead cross-section of the February 2010 fan chart

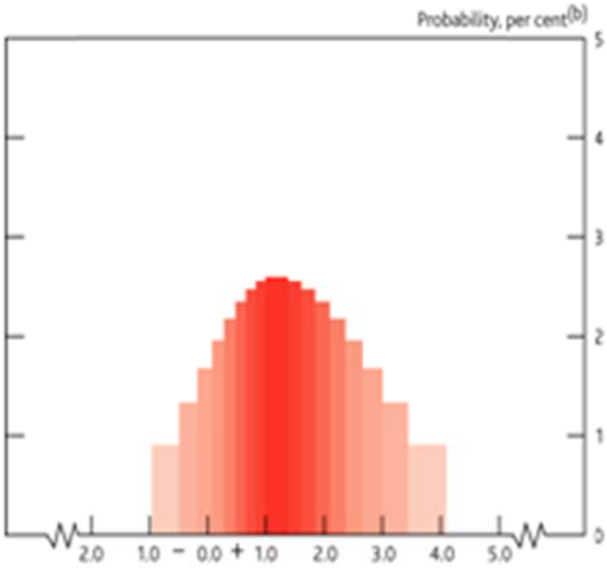


Figure 5. Spread of individual uncertainty measures, and median individual standard deviation
Upper panels: CPI inflation; lower panels: GDP growth; *h*=5, 9, 13

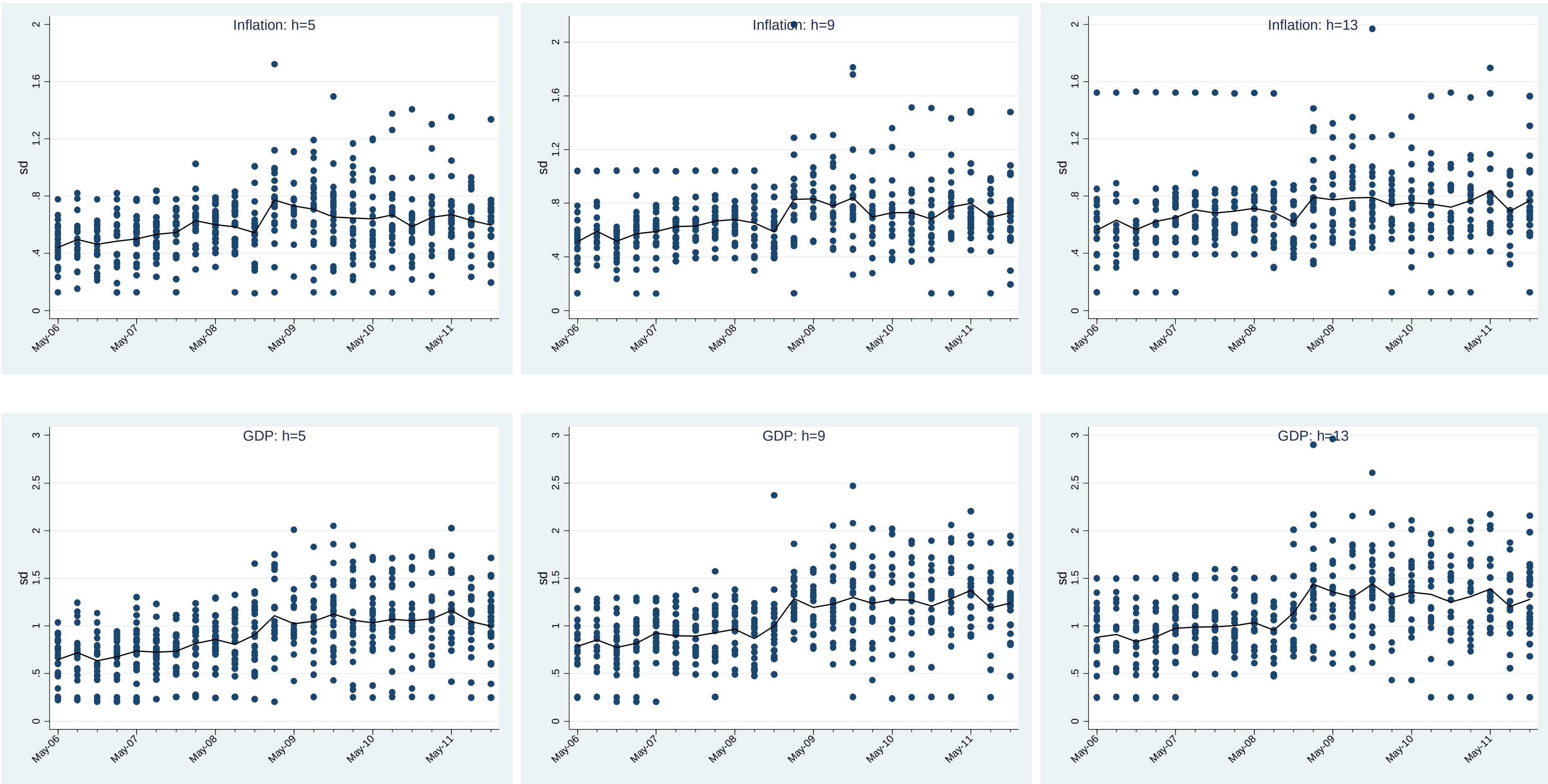


Figure 6. Uncertainty measures of the five highest-ranked (blue) and lowest-ranked (red) regular respondents in each panel
Upper panels: CPI inflation; lower panels: GDP growth; *h*=5, 9, 13

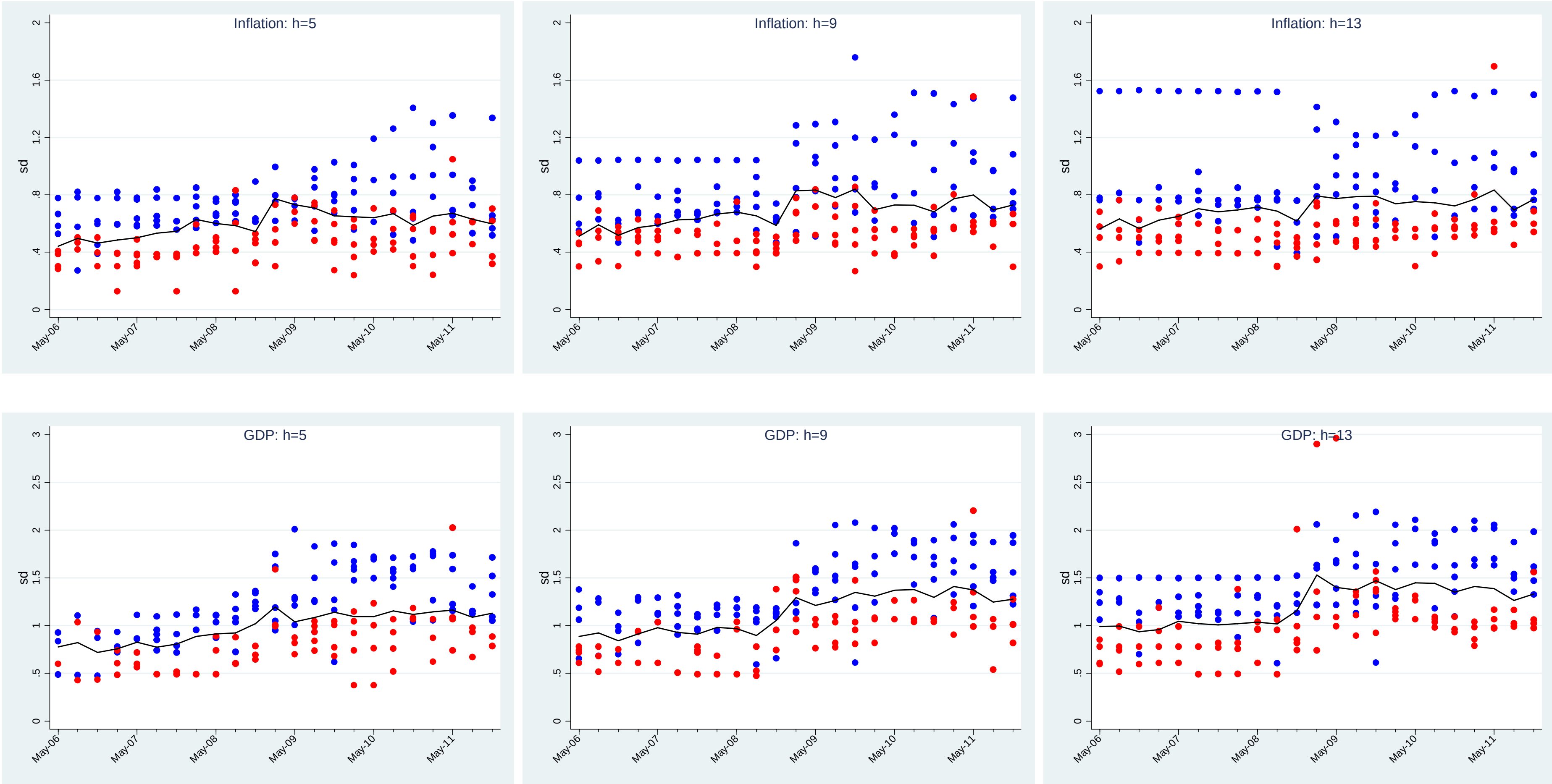


Figure 7. Spread of individual point forecasts, and survey mean point forecasts
Upper panels: CPI inflation; lower panels: GDP growth; *h*=5, 9, 13

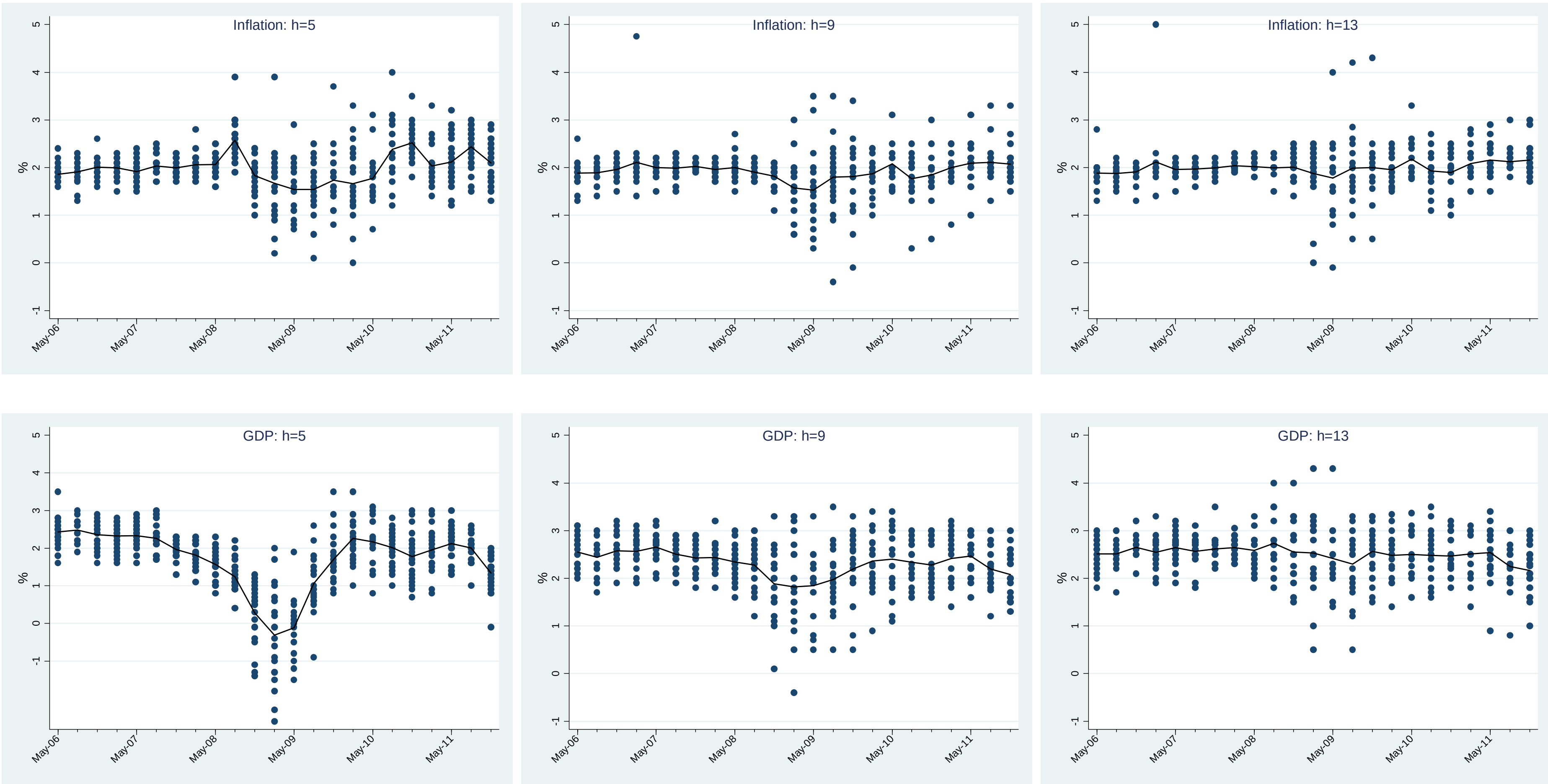


Figure 8. Point forecasts of the five highest-ranked (blue) and lowest-ranked (red) regular respondents in each panel
Upper panels: CPI inflation; lower panels: GDP growth; *h*=5, 9, 13

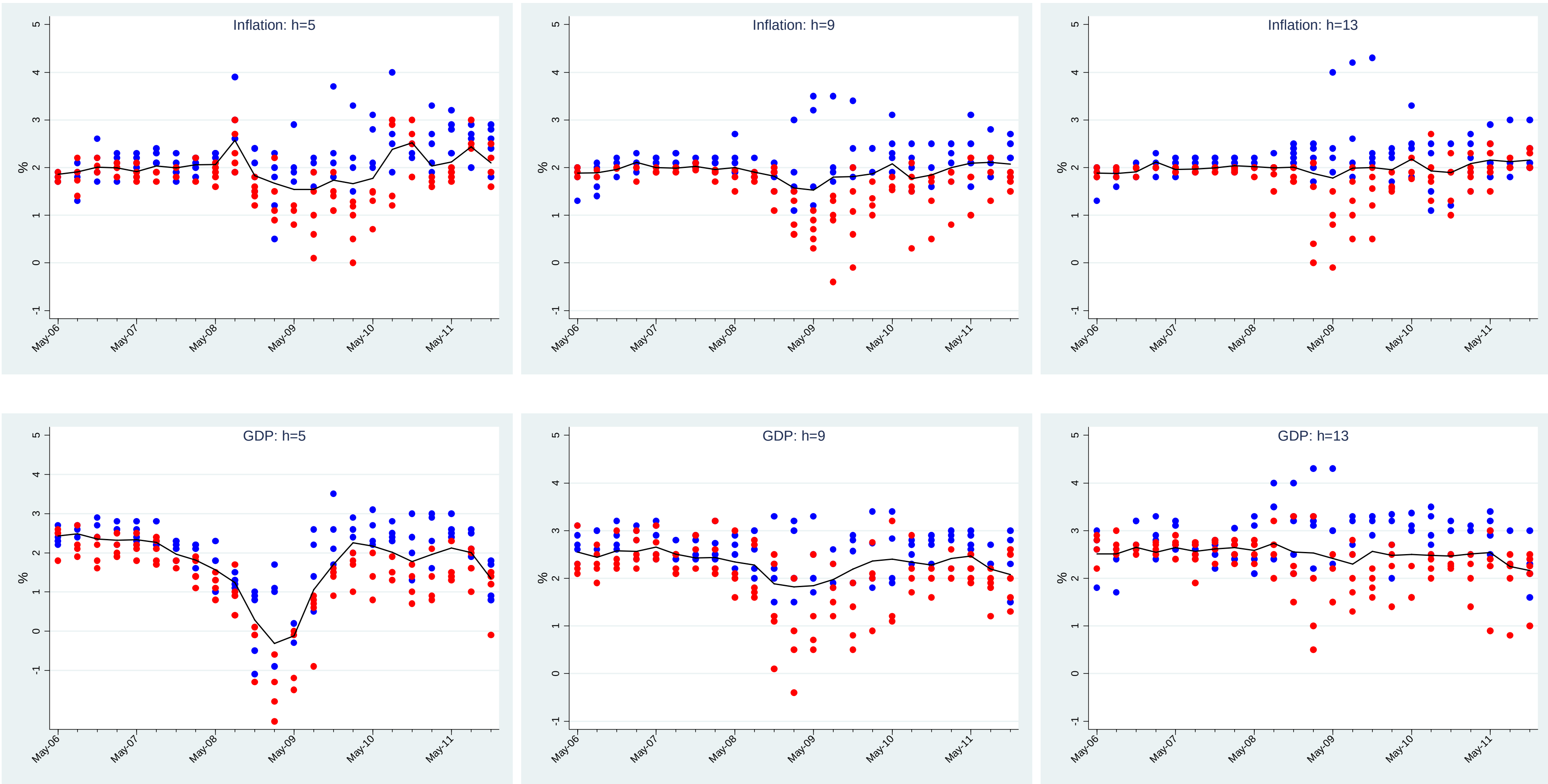


Figure 9. Aggregate variation, average individual uncertainty, and disagreement
Upper panels: CPI inflation; lower panels: GDP growth; *h*=5, 9, 13

Figure A1. ECB SPF mean point forecasts of inflation, and latest data available to forecasters

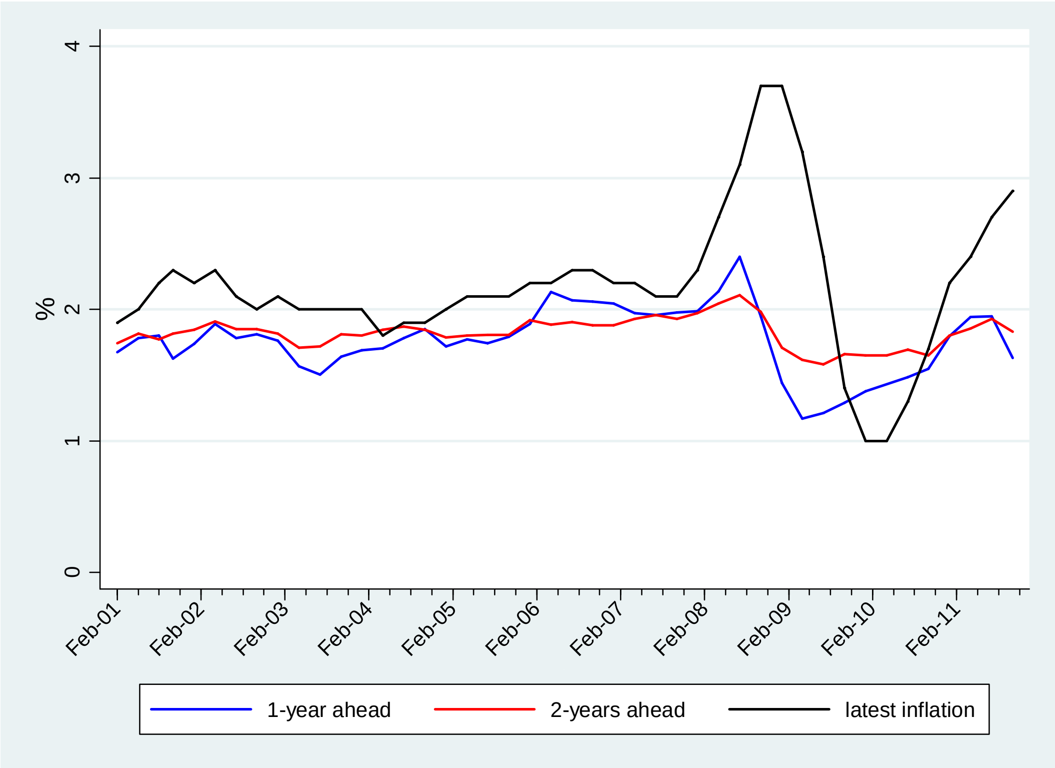


Figure A2. Spread of individual uncertainty measures, and median individual standard deviation, ECB SPF
Upper panels: all respondents; lower panels: eight highest-ranked (blue) and lowest-ranked (red) regular respondents